

Università degli Studi di Salerno  
Dipartimento di Scienze Aziendali - Management & Innovation System

Dottorato di Ricerca in Management and Information Technology  
Curriculum Informatica, Sistemi Informativi e Tecnologie del Software  
XVI Ciclo



Ph.D. Thesis

**Information Visualization:  
from Petroglyphs to CoDe Graphs**

Candidate

Dr. Paola de Roberto

Coordinator

Prof. Andrea De Lucia

Tutor

Prof. Genoveffa Tortora

2016 - 2017

# Abstract

Data visualization concerns the communication of data through visual representations and techniques. It aims at enhancing perception and support data-driven decision making so enabling insights otherwise hard to achieve. A good visualization of data makes it possible to identify patterns and enables better understanding of phenomena. In other words, data visualization is related to an innate human ability to quickly comprehend, discern and convert patterns into useful and usable information.

Humans have used visual graphical representations as early as 35.000 B.C., through cave drawings. Indeed, human ancestors already reasoned in terms of models or schemata: the visual representation of information is an ancient concept, as witnessed by the rock carvings found. Over the centuries, information visualization has evolved to take into account the changing human needs and its use has become more and more conscious. The first data visualization techniques have been developed to observe and represent physical quantities, geography and celestial positions. Successively, the combined use of euclidean geometry and algebra improved accuracy and complexity of information representation, in different fields, such as astronomy, physics and engineering. Finally, in the last century most modern forms of data representations were invented: starting from charts, histograms, and graphs up to high dimensional data, and dynamic and interactive visualizations of temporal data [41].

Nowadays, the huge amount of information enables more precise interpretation of phenomena so fostering the adoption of infographic techniques, in particular, for supporting managerial decision-making in the business area.

Actually, data visualization represents the heart of Business Analytics, which needs to

transform a large amount of complex data into comprehensible information for the most intuitive description of phenomena.

Information visualization uses visual reports which can be considered sentences of a proper visual language in a given specific domain. Then, there exists the need to examine and to address the main issues related to visual languages and information visualization in different domains: *i)* through the analysis and the interpretation of visual information; *ii)* by investigating the use of information modeling and the representation techniques needed to design and implement User-Centered decision support systems.

In the first part of the thesis, we propose a Visual Analytics system applied to the archaeological area; this system supports rupestrian archaeologists in the analysis and classification of thousands of reliefs and artifacts; each artifact contains different engravings, often called *petroglyphs*.

Rock art is a term coined in archeology for indicating any human made markings carved on natural stone [19,65]. The most part of the symbols concerning rock art are represented by *petroglyphs*, which were created by removing part of a rock surface by incising, picking, carving, and abrading. Although it is not possible to give certain interpretations to these petroglyphs, archaeologists have proposed many theories to explain their meaning, e.g., astronomical, cultural, or religious [105]. To this aim, we presents a visual analytics system, named *DARK*, for supporting rock art archaeologists in exploring repositories of rock art scenes each consisting of hundreds of petroglyphs carved by ancient people on rocks. With their increasing complexity, analyzing these repositories of heterogeneous information has become a major task and challenge for rock art archaeologists. *DARK* combines visualization techniques with fuzzy-based analysis of rock art scenes to infer information crucial for the correct interpretation of the scenes. Moreover, the *DARK* views allow archaeologists to validate their hypothesis against the information stored in the repository. In addition, we describe and detail the main features of the *PetroSketch* mobile application for supporting archaeologists in the classification and recognition of petroglyph symbols. *PetroSketch* is a virtual notebook enabling users to draw a petroglyph symbol on a white page, or by following the contour of a symbol captured with the camera, and to obtain its classification and the list of symbols more similar to it. The latter is performed by

a flexible image matching algorithm that measures the similarity between petroglyph by using a distance, derived from the image deformation model, which is computationally efficient and robust to local distortions.

In the second part of the thesis, we focus our attention on investigating the data visualization problem in the area of Business Analytics. In this context, the visual report of information extracted from data sources and in particular from data warehouses is faced. Visualization tools play a central role in contexts where information must be represented by preserving both the accuracy of data, and the complexity of relationships between data. Much attention has been paid to the problem of effective visualization of data in individual reports that usually are viewed through different types of standard graph (histogram, bar-plots, etc.) or in a tabular form. However, this kind of representation provides separate information items, but gives no support to visualize their relationships, which are the basis for most decision processes. Indeed, decision makers spend significant time and effort interpreting graphics derived from large multidimensional databases. Then, the choice of a graphical representation is critical whenever it is necessary to interpret data. Data are usually represented by several dashboard diagrams, such as histograms and pies, which do not highlight logical relationships among them.

The CoDe (Complexity Design) language allows to organize visualizations, named CoDe Graphs, by composing and aggregating dashboard diagrams through graphical conceptual links. The CoDe-based graph composition modeling allows to visualize relationships between information in the same image following the definition of efficiency of a visualization given by Jacques Bertin [9]: *"The most efficient (graphic) construction are those in which any question, whatever its type and level, can be answered in a single instant of perception, that is, in a single image"*. This representation named CoDe model can be considered a high-level cognitive map of the complex information underlying the ground data. The choice of the final visualization layout in terms of standard graphs is left to a visualization interface which provides the necessary implementation constructs.

The application of the CoDe methodology impacts principally on the Business Analytics field where the knowledge, management and analysis of company data (e.g., sales, production, costs and profits) are fundamental requirements for a valid decision-making



system. In fact, in modern companies, strategic information is necessary to grant the solidity and survival of business companies. In addition, CoDe is an alternative graphic model for the representation of data that reduces the gap between the data warehouse expert and company management. Indeed, CoDe allows to select the information of interest, to model the relationships between them and to automatically generate the diagrams containing this information linked through the relationships identified. To this aim, we presents the generation process of CoDe Graphs, and after analyzing the state of the art concerning the evaluation of graphical representation comprehensibility, we first propose a classification of that evaluation approaches and then perform the evaluation of the comprehensibility of CoDe Graphs with respect to dashboard reports by means of a controlled experiment, involving 47 participants. Results show that CoDe Graphs reduces participants' effort, while improving effectiveness and efficiency when a comprehension task is performed, so witnessing the usefulness of the CoDe methodology.